Chemical RNA Labeling without 3' End Bias Using Fluorescent *cis*-Platin Compounds

BioTechniques 34:974-980 (May 2003)

ABSTRACT

Recently, fluorescent, monofunctional *cis*-platin derivatives have been developed to chemically label nucleic acids for use in fluorescent hybridization assays. Here we show by hybridizations to microarrays containing oligonucleotide probes for the 3' ends, middle parts, and 5' ends of mRNAs, that this labeling methodology bypasses the problem of the 3' end bias that is characteristic of the conventional enzymatic oligo(dT)-primed, reverse transcription labeling of mRNAs.

INTRODUCTION

The advent of microarray and chip hybridization technology has enabled large-scale gene expression analysis at the RNA level by hybridizing fluorescent cDNA of an mRNA sample to large numbers of gene probes and taking the resulting fluorescence hybridization signals as measures of mRNA concentrations (1,2). In the great majority of such gene expression profiling studies, oligo(dT)-primed reverse transcription of poly(A) mRNA is applied to label the hybridization target. In a substantial fraction of these studies, the oligo(dT) primer is extended at the 5' end with a sequence that promotes cRNA transcription by a bacterial phage RNA polymerase, thus permitting, after second-strand cDNA synthesis, a linear target amplification (3) that compensates for limited sample size and sensitivity (4).

A well-established consequence of oligo(dT)-primed reverse transcription is that due to the limited processivity of the reverse transcriptase and secondary RNA structure, many of the mRNA species are represented in the labeled target only with their 3' ends, a bias that will also be reflected in amplified cRNA.

Here we present a chemical labeling

procedure that overcomes the 5' end representation problem. The method is based on fluorescent *cis*-platin derivatives that bind to DNA or RNA monofunctionally, with preference for the N7 of the guanosine residues, and has been applied in fluorescence in situ hybridization applications (5). Their use in cDNA microarray hybridizations has recently been described (6). These fluorescent platinum compounds label RNAs over their full lengths, as evidenced here by the efficient detection of hybridization signals from mRNA sequences hybridized to oligonucleotide arrays with the 3' ends, middle parts, and 5' ends represented as 70-mer oligonucleotides. The potential applications of this chemical RNA labeling methodology are briefly discussed.

MATERIALS AND METHODS

Genes, Oligonucleotides, and Microarrays

GenBank® accession numbers, the sequences, and positions of the (sense and antisense) oligonucleotides as distances from the 3' end are available upon request. Oligo(dT) sequences (70-mers; T_m range, 74°C–83°C) were synthesized by Qiagen/Operon Technologies (Venlo, The Netherlands) and

spotted in replicates of six on amino-silane-coated slides at 40 μ M using a GMS 417 arrayer (Affymetrix, Woburn, MA, USA).

Enzymatic cDNA Labeling, Chemical mRNA Labeling, and Hybridization

Enzymatically labeled cDNA was prepared from total Jurkat RNA using the MICROMAX™ Direct Labeling Kit (Perkin-Elmer Life Sciences, Boston, MA, USA) with CyTM3 and CyTM5 dUTP. Microarray hybridizations with the combined Cy3- and Cy5-labeled cDNAs were carried out for 16 h at 55°C with 10- μ L volumes in the hybridization buffer supplied with the kit with the equivalent of 50 μ g total RNA.

For chemical mRNA labeling, poly(A) mRNA was isolated using the OligoTex kit (Qiagen), followed by direct chemical labeling using 3 μ g/20 μ L mRNA in 5 mM Tris-HCl, pH 7.5, and Cy3 and Cy5 MICROMAX ASAP Labeling Reagents (Perkin-Elmer Life Sciences). Labeling was carried out at 85°C for 15 min. After the quenching of the reactions with 1% sodium diethyldithiocarbamate in 250 mM Tris-HCl, pH 7.5, the labeled mRNAs were combined and purified from quenched labeling reagents with YM100 spin columns (Millipore, Bedford, MA, USA). Following SpeedVac® drying, the labeled

Figure 1. ANOVA analysis of RNA representation of all 38 genes. Means of fluorescence intensities (representations) of all 38 genes are plotted relative to their oligonucleotide position (3' end, middle, or 5' end) and orientation (sense or antisense). The mean signal at the 5' position with enzymatic cDNA labeling differed significantly from the one at the 3' position ($P = 0.08$). Bars indicate standard errors. They are large primarily through the large differences in expression levels. Squares, chemical mRNA labeling; triangles, enzymatic cDNA labeling; open symbols, antisense probes; and closed symbols, sense probes.

Short Technical Reports

Figure 2. Correlation between Cy3 and Cy5 intensities in self-to-self hybridization of chemically labeled Jurkat mRNA. Data points are from signals of the 28 fully present genes.

mRNA targets were resuspended in 40 μ L hybridization buffer II of the MICROMAX ASAP kit. Microarray hybridizations were for 16 h at 50°C with 10- μ L volumes in Corning hybridization chambers (Corning, Corning, NY, USA) with the equivalent of 75 μ g total RNA.

Data Acquisition, Analysis of Variance, and Presence/Absence Criteria

Microarrays were scanned with an Axon 4000 scanner (Axon Instruments, Union City, CA USA) with the photomultiplier tube set at 650 V for Cy3 and Cy5 enzymatic cDNA labeling and 590 and 620 V for Cy5 and Cy3 chemical mRNA labeling. The fluorescence intensities of the spots were determined with the aid of ArrayPro (MediaCybernetics, Carlsbad, CA, USA).

An analysis of variance (ANOVA) of fluorescence intensities was performed for all genes collectively at the relative positions using SPSS software (Chicago, IL, USA).

For the analysis of individual genes, we defined the presence of a signal for mRNA direct chemical labeling at a given position in a replicate if, for the average of the six spots, the antisense value minus (sense value, plus $3 \times \text{SD}$) were greater than zero. For cDNA labeling, sense values minus (antisense value, plus $3 \times \text{SD}$) greater than zero were taken as present calls. For the presence of the full-length transcript, we demanded that at all three positions, a minimum of two of the three replicate microarray hybridizations gave present calls at all three positions in both colors; that is, a minimum of 12 calls of the maximum of 18 (equals three replicate hybridizations $\times$ two colors $\times$ three positions).

RESULTS

To analyze the capacity of the chemical mRNA labeling and the enzymatic cDNA labeling to report various segments of gene transcripts equally, we produced oligonucleotide “representation microarrays” for 38 gene transcripts (size range, 1259–12 783 nucleotides) by synthesizing sense and antisense oligonucleotide probe sequences of 70 nucleotides in length

Short Technical Reports

representing the 3' end, the middle, and the 5' end of the transcripts. The oligonucleotides were printed in replicates of six on aminosilane-activated microscopic glass slides. Because of rapid annealing of labeled single-stranded mRNA and cDNA targets, we opted to analyze results of separate, replicated microarray hybridizations with the different labeling methods in a self-to-self bicolor format. Thus, we hybridized three of the oligonucleotide microarrays with Cy3 and Cy5 chemically labeled Jurkat mRNA and three with enzymatically Cy3- and Cy5-dUTP labeled Jurkat cDNA [oligo(dT)-primed reverse transcription] and analyzed representation of the three positions globally for all genes as well as each gene individually.

ANOVA was used for global analysis (Figure 1) and showed the means of the representations at all three relative positions to be near equal for mRNA labeled with the platin reagents as reported by antisense probes. For cDNA labeling, the means of representation of all genes as reported on sense probes were significantly reduced as one moved to the 5' end, clearly illustrating the 3' end bias of the latter procedure. ANOVA allows for assessment of the factors that contribute to the variance observed. As expected from the widely varying levels of expression, the largest factor contributing to the variance was from the genes themselves (72%). The replication of the microarray hybridizations contributed 1.8% to the variance. The cyanine dyes also contribute to the variance in the same order of magnitude (1.6%), which was also reflected in excellent correlation between the Cy3 and Cy5 signals with the chemical RNA labeling (Figure 2; correlation coefficient in linear regression: 0.98). A portion (25.4%) of the variance remained unexplained (i.e., resulted from experimental variability of unknown source).

While the global ANOVA analyses may show the absence of positional bias with chemical mRNA labeling, it conceals how individual genes vary in representation at the different positions. To assess the equivalence of representation at the three positions of individual genes in a quantitatively reliable way, a stringent criterion of presence

Table 1. Chemical mRNA Labeling Gives Better 5' Representation than Enzymatic cDNA Labeling

Labeling Methodology	Relative Position of Oligonucleotide Probe on Transcript		
	3' End	Middle	5' End
mRNA (chemical)	28	28	28 (1133–8942)
cDNA (enzymatic)	24	17	12 (1133–2851)

The number of genes with signals present at the 3' end, middle, and 5' end positions are indicated. Numbers in parentheses indicate the absolute distance ranges (in nucleotides) of the 5' probes from the 3' end.

for a “full” transcript was set to eliminate analytically unreliable data from low expressing genes (see Material and Methods).

Twenty-eight genes fulfilled the criterion of presence for a full transcript for the chemical mRNA labeling (range

of distances from 3' to 5' end: 1133–8942 nucleotides; Figure 3A and Table 1). Of these 28, there were only 12 gene transcripts fully present in labeling by direct Cy-dUTP reverse transcription (42%; range of distances from 3' to 5' end: 1133–2851 nucleotides; Figure 3B

Figure 3. Fluorescence intensity as a function of probe position. (A) Twenty-eight gene transcripts fully present in chemically labeled mRNA. (B) Twelve gene transcripts fully present in enzymatically labeled cDNA. Two outliers are not shown in panel A.

Short Technical Reports

and Table 1). The reverse situation of gene transcripts fully present in cDNA but not in mRNA, labeling was not encountered.

Of the 10 gene transcripts that did not meet the criterion in mRNA labeling, four gave in one of the replicates clear, present calls at about equal intensity at all three positions. Among these four was the longest transcript selected (gene ID: U34994; 12 783 nucleotides).

When using a greater than 2-fold change as a commonly accepted value for the altered expression in RNA expression profiling by microarray hybridization, 20 (71%) transcripts could be considered equally represented at the three interrogated positions. When a 40% deviation was taken as a criterion of equal representation, 14 (50%) of the 28 gene transcripts could be considered as being expressed equally over their full lengths with the chemical mRNA labeling procedure. Five transcripts had consistent greater than 2-fold downward deviation (i.e., 3' end > middle > 5' end). Such underrepresentation of 5' ends can be explained by the selective mRNA degradation during the sample processing, which included oligo(dT) selection performed before the mRNA labeling. Three genes had a greater than 2-fold upward change. One of these had consistent upward deflection (3' end < middle < 5' end), one deviates about 8-fold at the 5' position only, and one deviates about 10-fold at the middle position only. BLAST searches did not indicate extensive sequence homologies for the latter two, and we consider these results as microarray hybridization artifacts. In enzymatic cDNA labeling, there was only one gene among the 12 fully present that complied with the less than 2-fold change criterion.

DISCUSSION

It is evident from the results that direct chemical mRNA labeling using fluorescent *cis*-platin compounds removes to a very large extent the 3' end bias that is characteristic of oligo(dT)-primed cDNA labeling of poly(A) RNA. As such, it may compete with enzymatic cDNA labeling using random primed reverse transcription,

which is also useful for detecting 5' RNA sequences (7).

Enzymatic cDNA and chemical RNA labeling both need relatively large amounts of input RNA. Enzymatic cRNA amplification with the incorporation of a label is frequently used to compensate for limited sample size or sensitivity (3,4). cRNA amplification, in the absence of labeled nucleotides, followed by the chemical platin labeling of the "cold" cRNA may alternatively be used, but the representation is then determined by the processivity of the reverse transcription during cDNA synthesis, and the specific advantage of transcriptome labeling with no positional bias may not be fully exploited.

The ability to faithfully label a transcriptome, together with its convenience, may make direct chemical *cis*-platin-based mRNA labeling a valuable technique for cDNA and cRNA microarray hybridizations in general (6) and the technique of choice for scanning genomic regions for expressed exons using tiled genomic oligonucleotide sequences (7) for the detection of alternative splice variants using oligonucleotides that span splice junctions (8) and for labeling non-poly(A) RNAs.

REFERENCES

1. DeRisi, J.L., V.R. Iyer, and P.O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
2. Schena, M., D. Shalon, R.W. Davis, and P.O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470.
3. Van Gelder, R.N., M.E. von Zastrow, A. Yool, W.C. Dement, J.D. Barchas, and J.H. Eberwine. 1990. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. USA* 87:1663-1667.
4. Duggan, D.J., M. Bittner, Y. Chen, P. Meltzer, and J.M. Trent. 1999. Expression profiling using cDNA microarrays. *Nat. Genet.* 21:10-14.
5. Wiegant, J.C., R.P. van Gijlswijk, R.J. Heetebrij, V. Bezrookove, A.K. Raap, and H.J. Tanke. 1999. ULS: a versatile method of labeling nucleic acids for FISH based on a monofunctional reaction of cisplatin derivatives with guanine moieties. *Cytogenet. Cell Genet.* 87:47-52.
6. Gupta, V., A. Cherkassky, P. Chatis, R. Joseph, A.L. Johnson, J. Broadbent, T. Erickson, and J. DiMeo. 2003. Directly labeled mRNA produces highly precise and unbiased differential gene expression data. *Nucleic Acids Res.* 31:e13.
7. Shoemaker, D.D., E.E. Schadt, C.D. Armour, Y.D. He, P. Garrett-Engele, P.D. McDonagh, P.M. Loerch, A. Leonardson, et al. 2001. Experimental annotation of the human genome using microarray technology. *Nature* 409:922-927.
8. Modrek, B. and C. Lee. 2002. A genomic view of alternative splicing. *Nat. Genet.* 30:13-19.

This work was supported in part by the Netherlands Organisation for Scientific Research NWO grant nos. 901-01-208 and 925-01-008. Address correspondence to Dr. Anton K. Raap, Department of Molecular Cell Biology, Leiden University Medical Center, 2333 AL Leiden, The Netherlands. e-mail: a.k.raap@lumc.nl

Received 28 January 2003; accepted 28 February 2003.

**R. Hagedoorn, R. Joseph¹,
S. Kasanmoentalib, P. Eilers,
J. Killian¹, and A.K. Raap**
*Leiden University Medical
Center
Leiden, The Netherlands
¹Perkin-Elmer Life Sciences
Boston, MA, USA*